

WHAT IS CLAIMED IS:

1. A method of describing the hierarchical structure of a document having content of a characteristic type of content, comprising:
5 expressing, independently of the document content type, the hierarchical structure of the document as a tree structure of one or more nodes; and
 providing a semantic representation for interpreting the tree structure.
2. The method of claim 1 further comprising associating with each tree node an
10 attribute describing the semantic character of the associated tree node.
3. The method of claim 1 wherein the semantic representation is provided based upon the document content type.
4. The method of claim 1 wherein the semantic representation is provided
15 independently of the document content.
5. The method of claim 1 wherein the semantic representation defines parent-child relationships among the nodes.
20
6. The method of claim 5 wherein expressing comprises associating with a node a child-count attribute indicative of whether the node has associated child nodes that have not yet been expressed in the tree structure.
7. The method of claim 1 further comprising:
25 expressing, independently of the document content type, the hierarchical structure of the document as a second tree structure of one or more nodes; and
 interpreting the second tree structure in accordance with a second semantic representation which is different from the first semantic representation.
8. The method of claim 1 wherein expressing comprises associating with a given
30 node an attribute identifying a second semantic interpretation for the structural feature

of the document represented by the given node, the second semantic interpretation being different from the first semantic interpretation.

5 9. The method of claim 1 further comprising recording the hierarchical structure of the document on a computer-readable medium.

10 10 A method of extracting content from a document having content of a characteristic type of content, comprising:

providing access to document content in response to a request for document content based upon an expression of the hierarchical structure of the document that is independent of document content type and has an interpretation controlled by a semantic representation.

15 11. The method of claim 10 further comprising providing the requested document content.

12 The method of claim 10 further comprising providing a pointer to the requested document content.

20 13. The method of claim 10 wherein the access to document content is provided in response to a client request.

14. A method of describing the hierarchical structure of a document having content of a characteristic type of content, comprising:

25 in response to a client request for information relating to the hierarchical structure of the document, expressing, independently of the document content type, the hierarchical structure of the document as a tree structure of one or more nodes; and providing a semantic representation for interpreting the tree structure.

30 15. The method of claim 14 wherein the client request comprises a request for information relating to the position of one or more nodes within the tree structure.

16. The method of claim 14 wherein expressing comprises expressing the hierarchical structure of the document at a level of detail specified in the client request.

5 17. The method of claim 14 wherein expressing comprises associating with a given node an attribute indicating the relative detail level represented by the given node.

10 18. The method of claim 17 further comprising:
in response to a client request for structural information about the given node at a level of detail that is different from the level of detail indicated by the attribute associated with the given node, expressing, independently of document content type, the hierarchical structure of the document as a tree structure of one or more nodes, including the given node, at the detail level specified in the client request.

15 19. The method of claim 14 wherein expressing comprises associating with a given node an attribute identifying a second semantic representation for the structural feature of the document represented by the given node, the second semantic interpretation being different from the first semantic interpretation.

20 20. The method of claim 19 further comprising, in response to a client request, providing the second semantic representation for interpreting the given node.

25 21. The method of claim 20 further comprising, in response to a client request, providing access to document content based on the second semantic representation.

22. A document description file, stored on a computer-readable medium, for describing the hierarchical structure of a document having content of a characteristic type of content, comprising:

30 a tree structure of one or more nodes expressing, independently of the document content type, the hierarchical structure of the document; and
a semantic representation for interpreting the tree structure.

23. The document description file of claim 22 further comprising an attribute associated with each tree node describing the semantic character of the associated tree node.

5 24. The document description file of claim 22 wherein the semantic representation is based upon the document content type.

25. The document description file of claim 22 wherein the semantic representation is independent of the document content.

10 26. The document description file of claim 22 wherein the semantic representation defines parent-child relationships among the nodes.

27. The document description file of claim 26 further comprising a child-count attribute associated with a node that is indicative of whether the node has associated child nodes that are not yet expressed in the tree structure.

28. The document description file of claim 22 further comprising:
a second document description file comprising
20 a second tree structure of one or more nodes expressing, independently of the document content type, the hierarchical structure of the document, and
a second semantic representation which is different from the first semantic representation for interpreting the second tree structure.

25 29. The document description file of claim 22 further comprising an attribute associated with a given node identifying a second semantic interpretation for the structural feature of the document represented by the given node, the second semantic interpretation being different from the first semantic interpretation.

30 30. A document description file, stored on a computer-readable medium, for describing the hierarchical structure of a document having content of a characteristic type of content, comprising:

a tree structure of one or more nodes expressing, independently of the document content type, the hierarchical structure of the document;

a semantic representation for interpreting the tree structure; and

information relating to document content within the hierarchical structure expressed by one or more tree nodes produced in response to a client request for document content associated with one or more tree nodes.

31. The document description file of claim 30 wherein the information relating to document content comprises a pointer to the requested document content.

32. The document description file of claim 30 wherein the information relating to document content comprises the requested document content.

33. A document description file, stored on a computer-readable medium, for describing the hierarchical structure of a document having content of a characteristic type of content, comprising:

a tree structure of one or more nodes expressing, independently of the document content type, the hierarchical structure of the document, the tree structure being produced in response to a client request for information relating to the hierarchical structure of the document; and

a semantic representation for interpreting the tree structure.

34. The document description file of claim 33 wherein the tree structure expresses the hierarchical structure of the document at a level of detail specified in the client request.

35. The document description file of claim 33 further comprising an attribute associated with a given node indicating the relative detail level represented by the given node.

36. The document description file of claim 35 further comprising:
a tree structure of one or more nodes, including the given node, expressing,
independently of document content type, the hierarchical structure of the document at
a detail level specified in a client request for structural information about the given
node, the requested detail level being different from the level of detail indicated by the
attribute associated with the given node.

37. The document description file of claim 33 further comprising an attribute
associated with a given node identifying a second semantic interpretation for the
structural feature of the document represented by the given node, the second semantic
interpretation being different from the first semantic interpretation.

38. A method executed on a computer, comprising:
generating a first document description file for describing a document stored on a
computer-readable medium, comprising:
generating a description of an application that produced the document;
generating a description of a location from which the document can be
obtained; and
generating a description of an operation that can be performed on the
document.

39. The method of claim 38, wherein the description of the location comprises a
uniform resource locator.

40. The method of claim 39, wherein the uniform resource locator identifies a
server configured to produce the document upon request.

41. The method of claim 39, wherein the uniform resource locator identifies a
location at which the document is stored.

42. The method of claim 38, wherein:

the operation comprises a transformation of the document from a file stored in a first storage format to a file stored in a second storage format; and

the operation produces a second document description file that describes the file stored in the second storage format.

5

43. The method of claim 42, further wherein the second document description file describes the first document description file.

44. The method of claim 38, wherein:

10

the operation comprises extraction of information from the document; and

the operation produces a second document description file that describes the information extracted from the document.

15

45. The method of claim 44, further wherein the second document description file describes the first document description file.

46. The method of claim 44, wherein the information extracted from the document describes a range of pages of the document.

20

47. The method of claim 44, wherein the document represents a multi-layered graphical object, and the information extracted from the document describes a subset of the layers of the multi-layered graphical object.

25

48. The method of claim 38, further comprising generating application-specific data describing the document.

49. The method of claim 48, wherein application-specific data comprises a name of an application that produced the document.

30

50. The method of claim 48, wherein application-specific data comprises a version number of an application that produced the document.

51. The method of claim 38, further comprising:
generating a field containing information describing the document.

52. The method of claim 51, wherein the field is an HTTP header.

53. The method of claim 51, wherein the field describes a date on which the document was produced.

54. The method of claim 51, wherein the field describes a date on which the document was modified.

55. The method of claim 51, wherein the field describes a size of the document.

56. The method of claim 51, wherein the field describes content contained in the document.

57. The method of claim 38, wherein the content of the first document description file is represented in XML syntax.

58. A method for processing a request for information derived from a first document, the first document being stored on a computer-readable medium, the first document being described by a first document description file stored on a computer-readable medium:

receiving the request and the first document description file from a client;
retrieving the information derived from the first document;
generating a second document description file describing the information derived from the first document.

59. The method of claim 58, wherein retrieving the information derived from the first document comprises retrieving a second document, stored on a computer readable medium, containing the information derived from the first document.

60. The method of claim 58, wherein retrieving the information derived from the first document comprises performing the operation on the first document to produce the information derived from the first document.

5 61. The method of claim 58, wherein the information derived from the first document comprises a second document.

62. The method of claim 61, wherein the second document description file comprises a pointer to the second document.

10 63. The method of claim 62, wherein the pointer comprises a uniform resource locator.

15 64. The method of claim 61, further comprising transmitting the second document to the client.

65. The method of claim 58, wherein the information derived from the first document comprises the first document.

20 66. The method of claim 58, wherein the information derived from the first document comprises a pointer to the first document.

67. The method of claim 66, wherein the pointer comprises a uniform resource locator.

25 68. The method of claim 58, further comprising:
transmitting the second document description file to the client.

30 69. The method of claim 58, wherein the first document description file contains the request.

70. A document description file, stored on a computer-readable medium, for describing a document stored on a computer-readable medium, the document description file comprising:

- a description of an application program that produced the document;
- a description of a location from which the document can be obtained; and
- a description of an operation that can be performed on the document.

71. The document description file of claim 70, wherein:

the operation that can be performed on the document comprises a transformation of the document from a file stored in a first storage format to a file stored in a second storage format; and

the operation produces a second document description file that describes the file stored in the second storage format.

72. The document description file of claim 70, wherein:

the operation that can be performed on the document comprises extraction of information from the document; and

the operation produces a second document description file that describes the information extracted from the document.

73. The document description file of claim 70, further comprising a description of an operation to be performed on the document.

74. A method executed on a computer for retrieving information derived from a first document, the first document being stored on a computer-readable medium, the method comprising:

retrieving by a first client a first document description file, stored on a computer-readable medium, in response to a request from a second client, the first document description file describing an application that produced the first document, a location from which the first document can be obtained, and an operation that can be performed on the first document;

using the first document description file to retrieve the information derived from the first document; and

transmitting the information derived from the first document to the second client.

5 75. The method of claim 74, wherein using the first document description file to retrieve the information derived from the first document comprises retrieving a second document, stored on a computer readable medium, containing the information derived from the first document.

10 76. The method of claim 74, wherein using the first document description file to retrieve the information derived from the first document comprises performing the operation on the first document to produce the information derived from the first document.

15 77. The method of claim 74, wherein the information derived from the first document comprises a second document.

20 78. The method of claim 74, wherein the information derived from the first document comprises the first document.

25 79. A system comprising:

a first computer-readable medium having a first document produced by a first application;

a second computer-readable medium having a first document description file describing the first application, a location on the first computer-readable medium from which the first document can be obtained, and a description of an operation that can be performed on the first document; and

a server configured to produce information derived from the first document.

30 80. The system of claim 79, wherein the information derived from the first document comprises a second document.

81. The system of claim 80, wherein the information derived from the first document comprises a second document description file describing the second document.

5 82. The system of claim 81, wherein the second document description file further describes the first document description file.

83. The system of claim 79, wherein the first and second computer-readable media are the same computer-readable media.

10 84. The system of claim 79, wherein the first and second computer-readable media are different computer-readable media.

15 85. A method executed on a computer, comprising:
generating a composite document description file for describing a combination of a first document description file and a second document description file, the method comprising:

20 generating a description of the combination;
generating a description of the first document description file; and
generating a description of the second document description file.

86. The method of claim 85, wherein the combination is an aggregation of the first document description file and the second document description file.

25 87. The method of claim 85, wherein the combination is an alternation of the first document description file and the second document description file.

30 88. The method of claim 85, further comprising:
generating a description of an operation that can be performed on the first document description file and the second document description file.

89. A composite document description file, stored on a computer-readable medium, for describing a combination of a first document description file and a second document description file, the composite document description file comprising:

- a description of the combination;
- a description of the first document description file; and
- a description of the second document description file.

90. The document description file of claim 89, wherein the combination is an aggregation of the first document description file and the second document description file.

91. The document description file of claim 89, wherein the combination is an alternation of the first document description file and the second document description file.

92. The document description file of claim 89, further comprising a description of an operation that can be performed on the first document description file and the second document description file.